

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/084,691DATE: 06/05/98
TIME: 13:20:34

INPUT SET: S26411.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANTS: BUKH, J., MILLER, R.H. AND
6 PURCELL, R.H.
7
8 (ii) TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
9 AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
10 CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
11 AND THE USE OF REAGENTS DERIVED FROM THESE
12 SEQUENCES IN DIAGNOSTIC METHODS AND
13 VACCINES
14
15 (iii) NUMBER OF SEQUENCES: 263
16
17 (iv) CORRESPONDENCE ADDRESS:
18 (A) ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
19 (B) STREET: 345 PARK AVENUE
20 (C) CITY: NEW YORK
21 (D) STATE: NEW YORK
22 (E) COUNTRY: USA
23 (F) ZIP: 10154
24
25 (v) COMPUTER READABLE FORM:
26 (A) MEDIUM TYPE: FLOPPY DISK
27 (B) COMPUTER: IBM PC COMPATIBLE
28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
29 (D) SOFTWARE: WORDPERFECT 5.1
30
31 (vi) CURRENT APPLICATION DATA:
32 (A) APPLICATION NUMBER: TO BE ASSIGNED
33 (B) FILING DATE: 26-MAY-1998
34
35 (vii) PRIOR APPLICATION DATA:
36 (A) APPLICATION NUMBER: 08/290,665
37 (B) FILING DATE: 15-AUG-1994
38
39 (vii) PRIOR APPLICATION DATA:
40 (A) APPLICATION NUMBER: 08/086,428
41 (B) FILING DATE: 29-JUNE-1993
42
43 (viii) ATTORNEY/AGENT INFORMATION:
44 (A) NAME: RICHARD W. BORK
45 (B) REGISTRATION NUMBER: 36,459
46 (C) REFERENCE/DOCKET NUMBER: 2026-4116US2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/084,691DATE: 06/05/98
TIME: 13:20:35

INPUT SET: S26411.raw

47
48 (ix) TELECOMMUNICATION INFORMATION:
49 (A) TELEPHONE: (212) 758-4800
50 (B) TELEFAX: (212) 751-6849
51 (C) TELEX: 421792
52
53 (2) INFORMATION FOR SEQ ID NO:1:
54
55 (i) SEQUENCE CHARACTERISTICS:
56 (A) LENGTH: 576 base pairs
57 (B) TYPE: nucleic acid
58 (C) STRANDEDNESS: single
59 (D) TOPOLOGY: linear
60
61 (vi) ORIGINAL SOURCE:
62 (A) ORGANISM: homosapiens
63 (C) INDIVIDUAL ISOLATE: DK7
64
65 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
66
67 TAC CAA GTG CGC AAC TCC ACG GGG CTT TAC CAT GTC ACC 39
68 AAT GAT TGC CCT AAC TCG AGT ATC GTG TAC GAG GCG GCC 78
69 GAT GCC ATC CTG CAC ACT CCG GGG TGT GTC CCT TGC GTT 117
70 CGC GAG GGT AAC GTC TCG AGG TGT TGG GTG GCG ATG ACC 156
71 CCC ACG GTG GCC ACC AGG GAT GGC AAA CTC CCC ACA GCG 195
72 CAG CTT CGA CGT CAC ATC GAT CTG CTC GTC GGG AGT GCC 234
73 ACC CTC TGT TCG GCC CTC TAC GTG GGG GAC CTG TGC GGG 273
74 TCT GTC TTT CTT GTC GGT CAA CTG TTT ACC TTC TCT CCC 312
75 AGG CGC CAC TGG ACG ACG CAA GGC TGC AAT TGT TCT ATC 351
76 TAT CCT GGC CAT ATA ACG GGT CAC CGC ATG GCG TGG GAT 390
77 ATG ATG ATG AAC TGG TCC CCT ACC ACG GCG TTG GTA GTA 429
78 GCT CAG CTG CTC CGG ATC CCG CAA GCC ATC TTG GAC ATG 468
79 ATC GCT GGT GCT CAC TGG GGA GTC CTG GCG GGC ATA GCG 507
80 TAT TTT TCC ATG GTG GGG AAC TGG GCG AAG GTC CTG GTA 546
81 GTG CTG CTG CTA TTT GCC GGC GTC GAC GCG 576
82
83
84 (2) INFORMATION FOR SEQ ID NO:2:
85
86 (i) SEQUENCE CHARACTERISTICS:
87 (A) LENGTH: 576 base pairs
88 (B) TYPE: nucleic acid
89 (C) STRANDEDNESS: single
90 (D) TOPOLOGY: linear
91
92 (vi) ORIGINAL SOURCE:
93 (A) ORGANISM: homosapiens
94 (C) INDIVIDUAL ISOLATE: DK9
95
96 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
97
98 TAC CAA GTA CGC AAC TCC TCG GGC CTC TAC CAT GTC ACC 39
99 AAT GAT TGC CCT AAC TCG AGT ATT GTG TAC GAG GCG GCC 78

RAW SEQUENCE LISTING PATENT APPLICATION US/09/084,691

DATE: 06/05/98
TIME: 13:20:36

INPUT SET: S26411.raw

100	GAT GCC ATC CTG CAT TCT CCA GGG TGT GTC CCT TGC GTT	117
101	CGC GAG GGT AAC GCC TCG AAA TGT TGG GTG GCG GTG GCC	156
102	CCC ACG GTG GCC ACC AGG GAC GGC AAG CTC CCC GCA ACG	195
103	CAG CTT CGA CGT CAC ATC GAT CTG CTT GTC GGG AGC GCC	234
104	ACC CTC TGC TCG GCC CTC TAT GTG GGG GAC TTG TGC GGG	273
105	TCT GTC TTC CTT GTC GGC CAA CTG TTC ACC TTC TCC CCC	312
106	AGA CGC CAC TGG ACA ACG CAA GAC TGC AAC TGT TCT ATC	351
107	TAC CCC GGC CAT ATT ACG GGT CAT CGC ATG GCG TGG GAT	390
108	ATG ATG ATG AAC TGG TCC CCT ACA GCA GCG CTG GTA ATG	429
109	GCG CAG CTG CTC AGG ATC CCG CAG GCC ATC TTG GAC ATG	468
110	ATC GCT GGT GCC CAC TGG GGA GTC CTA GCG GGC ATA GCG	507
111	TAT TTC TCC ATG GTG GGG AAC TGG GCG AAG GTC GTG GTG	546
112	GTA CTG TTG CTG TTT ACC GGC GTC GAT GCG	576

113

114

115 (2) INFORMATION FOR SEQ ID NO:3:

116

117 (i) SEQUENCE CHARACTERISTICS:

118 (A) LENGTH: 576 base pairs

119 (B) TYPE: nucleic acid

120 (C) STRANDEDNESS: single

121 (D) TOPOLOGY: linear

122

123 (vi) ORIGINAL SOURCE:

124 (A) ORGANISM: homosapiens

125 (C) INDIVIDUAL ISOLATE: DR1

126

127 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

128

129	CAC CAA GTG CGC AAC TCT ACA GGG CTT TAC CAT GTC ACC	39
130	AAT GAT TGC CCT AAT TCG AGT ATT GTG TAC GAG GCG GCC	78
131	GAT GCC ATC CTG CAC GCG CCG GGG TGT GTC CCT TGC GTT	117
132	CGC GAG GGT AAC GCC TCG AGG TGT TGG GTG GCG GTG ACC	156
133	CCC ACG GTG GCC ACC AGG GAC GGC AAA CTC CCC ACA ACG	195
134	CAG CTT CGA CGT CAC ATC GAC CTG CTT GTC GGG AGC GCC	234
135	ACC CTC TGC TCG GCC CTC TAC GTG GGG GAC CTG TGC GGG	273
136	TCT GTC TTC CTT GTC GGT CAA CTG TTC ACC TTT TCT CCC	312
137	AGG CGC CAC TGG ACA ACG CAA GAC TGC AAT TGT TCT ATC	351
138	TAT CCC GGC CAT ATA ACG GGA CAC CGT ATG GCA TGG GAT	390
139	ATG ATG ATG AAC TGG TCC CCT ACG ACA GCG CTG GTA ATG	429
140	GCT CAG CTG CTC CGG ATC CCA CAA GCC ATC TTG GAC ATG	468
141	ATC GCT GGA GCC CAC TGG GGA GTC CTA GCG GGC ATA GCG	507
142	TAT TTC TCC ATG GTG GGG AAC TGG GCG AAG GTC GTG GTA	546
143	GTG CTG TTG CTG TTT GCC GGC GTT GAT GCG	576

144

145

146 (2) INFORMATION FOR SEQ ID NO:4:

147

148 (i) SEQUENCE CHARACTERISTICS:

149 (A) LENGTH: 576 base pairs

150 (B) TYPE: nucleic acid

151 (C) STRANDEDNESS: single

152 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/084,691DATE: 06/05/98
TIME: 13:20:36

INPUT SET: S26411.raw

```
153
154      (vi) ORIGINAL SOURCE:
155          (A) ORGANISM:  homosapiens
156          (C) INDIVIDUAL ISOLATE:  DR4
157
158      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
159
160      CAC CAA GTG CGC AAC TCT ACA GGG CTT TAC CAT GTC ACC      39
161      AAT GAT TGC CCT AAT TCG AGT ATT GTG TAC GAG GCG GCC      78
162      GAT GCC ATC CTG CAC ACG CCG GGG TGT GTC CCT TGC GTT      117
163      CGC GAG GGT AAC ACC TCG AGG TGT TGG GTG GCG GTG ACC      156
164      CCC ACG GTG GCC ACC AGG GAC GGC AAA CTC CCC ACA ACG      195
165      CAG CTC CGA CGT CAC ATC GAC CTG CTT GTC GGG AGC GCC      234
166      ACC CTC TGC TCG GCC CTC TAC GTG GGG GAC TTG TGC GGG      273
167      TCT GTC TTC CTT GTC GGT CAA CTG TTC ACC TTC TCT CCC      312
168      AGG CAC CAC TGG ACA ACG CAA GAC TGC AAT TGT TCC ATC      351
169      TAT CCC GGC CAT ATA ACG GGC CAC CGC ATG GCG TGG GAT      390
170      ATG ATG ATG AAC TGG TCC CCT ACG ACA GCG CTG GTA GTA      429
171      GCT CAG CTG CTC CGG ATC CCA CAA GCC ATC TTG GAC ATG      468
172      ATC GCT GGT GCC CAC TGG GGA GTC CTA GCG GGC ATA GCG      507
173      TAT TTC TCC ATG GTG GGG AAC TGG GCG AAG GTC CTG GTA      546
174      GTG CTG TTG CTG TTT GCC GGC GTT GAT GCG      576
175
176
177      (2) INFORMATION FOR SEQ ID NO:5:
178
179      (i) SEQUENCE CHARACTERISTICS:
180          (A) LENGTH:  576 base pairs
181          (B) TYPE:    nucleic acid
182          (C) STRANDEDNESS:  single
183          (D) TOPOLOGY:  linear
184
185      (vi) ORIGINAL SOURCE:
186          (A) ORGANISM:  homosapiens
187          (C) INDIVIDUAL ISOLATE:  S14
188
189      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
190
191      TAC CAA GTG CGC AAC TCC ACG GGG CTT TAC CAT GTT ACC      39
192      AAT GAT TGC CCT AAC TCG AGT ATT GTG TAC GAG ACA GCT      78
193      GAT GCT ATC CTA CAC GCT CCG GGA TGT GTC CCT TGC GTT      117
194      CGT GAG GGT AAC ACC TCG AGG TGT TGG GTG GCG ATG ACC      156
195      CCC ACG GTG GCC ACC AGG GAC GGC AAA CTC CCC GCA ACG      195
196      CAG CTT CGA CGT TAC ATC GAT CTG CTT GTC GGG AGC GCC      234
197      ACC CTC TGT TCG GCC CTC TAC GTG GGG GAC TTG TGC GGG      273
198      TCT GTC TTT CTT GTC GGT CAG CTG TTT ACC TTC TCT CCC      312
199      AGG CGC CTC TGG ACG ACG CAA GAC TGC AAT TGT TCT ATC      351
200      TAT CCC GGC CAT ATA ACG GGT CAT CGC ATG GCA TGG GAT      390
201      ATG ATG ATG AAC TGG TCC CCT ACG ACG GCA CTG GTA GTA      429
202      GCT CAG CTG CTC CGG ATC CCA CAA GCC ATC TTG GAT ATG      468
203      ATC GCT GGT GCT CAC TGG GGA GTC CTA GCG GGC ATA GCG      507
204      TAT TTC TCC ATG GTG GGA AAC TGG GCG AAG GTC CTA GTG      546
205      GTG CTG CTG CTA TTC GCC GGC GTT GAC GCG      576
```

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/084,691DATE: 06/05/98
TIME: 13:20:37

INPUT SET: S26411.raw

206

207

208 (2) INFORMATION FOR SEQ ID NO:6:

209

210 (i) SEQUENCE CHARACTERISTICS:

211 (A) LENGTH: 576 base pairs

212 (B) TYPE: nucleic acid

213 (C) STRANDEDNESS: single

214 (D) TOPOLOGY: linear

215

216 (vi) ORIGINAL SOURCE:

217 (A) ORGANISM: homosapiens

218 (C) INDIVIDUAL ISOLATE: S18

219

220 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

221

222	TAC CAA GTA CGC AAC TCC ACG GGC CTT TAC CAT GTC ACC	39
223	AAT GAC TGC CCT AAC TCG AGC ATT GTG TAC GAG ACG GCC	78
224	GAT ACC ATC CTA CAC TCT CCG GGG TGT GTC CCT TGC GTT	117
225	CGC GAG GGT AAC GCC TCG AGA TGT TGG GTG CCG GTG GCC	156
226	CCC ACA GTT GCC ACC AGG GAC GGC AAA CTC CCC GCA ACG	195
227	CAG CTT CGA CGT CAC ATC GAT CTG CTT GTT GGG AGC GCC	234
228	ACC CTC TGC TCG GCC CTC TAT GTG GGG GAC CTG TGC GGG	273
229	TCT GTC TTT CTT GTC AGC CAG CTG TTC ACT ATC TCC CCC	312
230	AGG CGC CAC TGG ACA ACG CAA GAC TGC AAC TGT TCT ATC	351
231	TAC CCC GGC CAT ATA ACG GGT CAC CGT ATG GCA TGG GAT	390
232	ATG ATG ATG AAC TGG TCC CCT ACA ACG GCG TTG GTA ATA	429
233	GCT CAG CTG CTC AGG GTC CCG CAA GCC GTC TTG GAC ATG	468
234	ATC GCT GGT GCC CAC TGG GGA GTC CTA GCG GGC ATA GCG	507
235	TAT TTC TCC ATG GCG GGG AAC TGG GCG AAG GTC CTG CTA	546
236	GTG CTG TTG CTG TTT GCC GGC GTC GAT GCG	576

237

238

239 (2) INFORMATION FOR SEQ ID NO:7:

240

241 (i) SEQUENCE CHARACTERISTICS:

242 (A) LENGTH: 576 base pairs

243 (B) TYPE: nucleic acid

244 (C) STRANDEDNESS: single

245 (D) TOPOLOGY: linear

246

247 (vi) ORIGINAL SOURCE:

248 (A) ORGANISM: homosapiens

249 (C) INDIVIDUAL ISOLATE: SW1

250

251 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

252

253	TAC CAA GTA CGC AAC TCC TCG GGC CTT TAC CAT GTC ACC	39
254	AAT GAT TGC CCT AAC TCG AGT ATT GTG TAC GAG ACG GCC	78
255	GAT GCC ATT CTA CAC TCT CCA GGG TGT GTC CCT TGC GTT	117
256	CGC GAG GAT GGC GCC CCG AAG TGT TGG GTG GCG GTG GCC	156
257	CCC ACA GTC GCC ACT AGG GAC GGC AAA CTC CCT GCA ACG	195
258	CAG CTT CGA CGT CAC ATC GAT CTG CTT GTC GGA AGC GCC	234

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/084,691

DATE: 06/05/98
TIME: 13:20:38

INPUT SET: S26411.raw

Line	Error	Original Text
32	Wrong application Serial Number	(A) APPLICATION NUMBER: TO BE ASSIGNED